

\*\*\*\*\*  
 WISEA (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Thu Dec 23 10:13:23 1999; MasPar time 2.04 Seconds  
 Tabular output not generated. 83.062 Million cell updates/sec

Title: >US-09-177-843-2  
 Description: (1-6) from US09177843.ppe  
 Perfect Score: 41  
 Sequence: 1 GRGESP 6  
 Scoring table: PAM 150  
 Gap 15  
 Searched: 77977 seqs, 28268293 residues  
 Post-processing: Minimum Match 0%  
 Listing first 45 summaries  
 Database: swiss-prot37  
 i:swissprot

Statistics: Mean 18.780; Variance 16.567; scale 1.134

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	40	97.6	1256	1	FINC_CHICK FIBRONECTIN (FN) (FRAG	1.94e+00
2	40	97.6	1328	1	FINC_PLEWA FIBRONECTIN (FN) (FRAG	1.94e+00
3	40	97.6	2265	1	FINC_BOVIN FIBRONECTIN (FN)	1.94e+00
4	40	97.6	2386	1	FINC_HUMAN FIBRONECTIN PRECURSOR	1.94e+00
5	40	97.6	2477	1	FINC_MOUSE FIBRONECTIN PRECURSOR	1.94e+00
6	40	97.6	2477	1	FINC_RAT FIBRONECTIN PRECURSOR	1.94e+00
7	40	97.6	2481	1	FINC_XENLA FIBRONECTIN PRECURSOR	1.94e+00
8	39	95.1	425	1	YIK3_YEAST HYPOTHETICAL 48.3 KD P	3.67e+00
9	39	95.1	463	1	Y863_SYN3 HYPOTHETICAL 50.4 KD P	3.67e+00
10	38	92.7	336	1	FILA_MOUSE FILAGGRIN (FRAGMENT)	6.87e+00
11	38	92.7	370	1	HUPK_RHILV HYDROGENASE EXPRESSION	6.87e+00
12	38	92.7	492	1	CPS3_PIG CYTOCHROME P450 XXIA3	6.87e+00
13	38	92.7	492	1	CPS3_PIG CYTOCHROME P450 XXIA3	6.87e+00
14	38	92.7	495	1	CPB1_BOVIN CYTOCHROME P450 2E1 (E	6.87e+00
15	38	92.7	495	1	YNN7_YEAST HYPOTHETICAL 109.8 KD	6.87e+00
16	37	90.2	825	1	SE5_RAT 5E5 ANTIGEN	1.27e+01
17	37	90.2	1266	1	NGCA_CHICK NEURONAL-GLIAL CELL AD	1.27e+01
18	36	87.8	165	1	RS10_XENLA 40S RIBOSOMAL PROTEIN	2.31e+01
19	36	87.8	235	1	NHAB_RHOSO NITRILE HYDRATASE SUBU	2.31e+01
20	36	87.8	490	1	CPCH_HUMAN CYTOCHROME P450 2C8 (E	2.31e+01
21	36	87.8	583	1	PUP6_CRYNE PHOSPHORIBOSYLAMINOIM	2.31e+01
22	36	87.8	657	1	MY16_MOUSE MYELOID DIFFERENTIATIO	2.31e+01
23	36	87.8	715	1	Y07J_MYCTU HYPOTHETICAL 78.2 KD P	2.31e+01

24	36	87.8	831	1	PRIA_SYN3 PRIMSOMAL PROTEIN N'	2.31e+01
25	36	87.8	1181	1	HAIR_RAT HAIRLESS PROTEIN	2.31e+01
26	35	85.4	165	1	RS10_RAT 40S RIBOSOMAL PROTEIN	4.16e+01
27	35	85.4	165	1	RS10_HUMAN 40S RIBOSOMAL PROTEIN	4.16e+01
28	35	85.4	294	1	POL_SMSAV POL POLYPROTEIN [CONTA	4.16e+01
29	35	85.4	350	1	RS40_ARATH ARGININE/SERINE-RICH S	4.16e+01
30	35	85.4	350	1	STRR_STRGR STREPTOMICIN BIOSYNTH	4.16e+01
31	35	85.4	356	1	RS41_ARATH ARGININE/SERINE-RICH S	4.16e+01
32	35	85.4	388	1	RECF_CAUCR RECF PROTEIN	4.16e+01
33	35	85.4	527	1	KITH_HSVSA THYMIDINE KINASE (EC 2	4.16e+01
34	35	85.4	567	1	GPV_MOUSE PLATELET GLYCOPROTEIN	4.16e+01
35	35	85.4	567	1	GPV_RAT PLATELET GLYCOPROTEIN	4.16e+01
36	35	85.4	704	1	GLGB_YEAST 1,4-ALPHA-GLUCAN BRANC	4.16e+01
37	35	85.4	737	1	RECG_MYCTU ATP-DEPENDENT DNA HELI	4.16e+01
38	35	85.4	916	1	TOP1_ARATH DNA TOPOISOMERASE I (E	4.16e+01
39	35	85.4	972	1	TOP1_DROME DNA TOPOISOMERASE I (E	4.16e+01
40	35	85.4	1065	1	VINC_MOUSE VINCULIN	4.16e+01
41	35	85.4	1065	1	VINC_HUMAN VINCULIN	4.16e+01
42	35	85.4	1065	1	VINC_CHICK VINCULIN	4.16e+01
43	35	85.4	2411	1	MYSA_DROME MYOSIN HEAVY CHAIN, MU	4.16e+01
44	34	82.9	105	1	YEHK_ECOLI HYPOTHETICAL 12.6 KD P	7.37e+01
45	34	82.9	528	1	YRU4_CAEEL HYPOTHETICAL 59.7 KD P	7.37e+01

## ALIGNMENTS

RESULT 1  
 ID FINC\_CHICK STANDARD; PRT; 1256 AA.  
 AC P11722: 090921;  
 DT 01-OCT-1989 (REL. 12, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE FIBRONECTIN (FN) (FRAGMENTS).  
 GN FNL.  
 OS GALLUS GALLUS (CHICKEN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;  
 OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.  
 RN [1]  
 RP SEQUENCE OF 1-50 FROM N.A.  
 RX MEDLINE; 83117850.  
 RA HIRANO H., YAMADA Y., SULLIVAN M., DE CROMBRUGGHE B., PASTAN I.,  
 RA YAMADA K.M.;  
 RT "Isolation of genomic DNA clones spanning the entire fibronectin  
 RT gene.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 80:46-50(1983).  
 RN [2]  
 RP SEQUENCE OF 51-1256 FROM N.A.  
 RX STRAIN-WHITE LEHORN;  
 RA NORTON P.A.;  
 RL SUBMITTED (FEB-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [3]  
 RP SEQUENCE OF 227-415 FROM N.A.  
 RX MEDLINE; 96183658.  
 RA GERRIS A.L., BRANDLI D.W., LEWIS S.D., BENNETT V.D.;  
 RT "The exon encoding the fibronectin type III-9 repeat is  
 RT constitutively included in the mRNA from chick limb mesenchyme and  
 RT cartilage.";  
 RL BIOCHIM. BIOPHYS. ACTA 1311:5-12(1996).  
 RN [4]  
 RP SEQUENCE OF 327-599 FROM N.A.  
 RX MEDLINE; 88050950.  
 RA KUBOMURA S., OBARA M., KARASAKI Y., TANIGUCHI H., GOTOH S.,  
 RA TSUDA T., HIGASHI K., OHSATO K., HIARNO H.;  
 RT "Genetic analysis of the cell binding domain region of the chicken  
 RT fibronectin gene.";  
 RL BIOCHIM. BIOPHYS. ACTA 910:171-181(1987).  
 RN [5]  
 RP SEQUENCE OF 413-1256 FROM N.A.  
 RX MEDLINE; 88142820.  
 RA NORTON P.A., HYNES R.O.;  
 RT "Alternative splicing of chicken fibronectin in embryos and in normal  
 RT and transformed cells.";  
 RL MOL. CELL. BIOL. 7:4297-4307(1987).

```
Query Match      97.6%; Score 40; DB 1; Length 1256;
Best Local Similarity 83.3%; Pred. No. 1.94e+00;
Matches      5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

5

```

FT DISULFID 1324 1324 INTERCHAIN (WITH 1320 OF OTHER CHAIN) (BY
FT SITE 461 463 SIMILARITY).
FT CARBOHYD 89 89 CELL ATTACHMENT SITE.
SQ SEQUENCE 1328 AA; 145037 MW; 1139F7B6 CRC32;

Query Match 97.6%; Score 40; DB 1; Length 1328;
Best Local Similarity 83.3%; Pred. No. 1.94e+00;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 460 GRGSP 465
|||:|
1 GRGSP 6

RESULT 3
ID FINE_BOVIN STANDARD; PRT: 2265 AA.
AC P07589;
DT 01-APR-1988 (REL. 07, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE FIBRONECTIN (FN).
GN FNI.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; META20A; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.
RN [1]
RP SEQUENCE.
RX MEDLINE; 87054047.
RA SKORSTENGAARD K., JENSEN M.S., SAHL P., PETERSEN T.E., MAGNUSSON S.;
RT "Complete primary structure of bovine plasma fibronectin.";
RL EUR. J. BIOCHEM. 161:441-453(1986).
RN [2]
RP PARTIAL SEQUENCE.
RX MEDLINE; 83117805.
RA PETERSEN T.E., THORGERSEN H.C., SKORSTENGAARD K., VIBE-PEDERSEN K.,
RA SAHL P., SOTTRUP-JENSENS L., MAGNUSSON S.;
RT "Partial primary structure of bovine plasma fibronectin: three types
of internal homology.";
RL PROC. NATL. ACAD. SCI. U.S.A. 80:137-141(1983).
RN [3]
RP SEQUENCE OF 2170-2265 FROM N.A.
RX MEDLINE; 83221567.
RA KORNBLIHT A.R., VIBE-PEDERSEN K., BARALLE F.E.;
RT "Isolation and characterization of cDNA clones for human and bovine
fibronectins.";
RL PROC. NATL. ACAD. SCI. U.S.A. 80:3218-3222(1983).
CC -!- FUNCTION: FIBRONECTINS BIND CELL SURFACES AND VARIOUS COMPOUNDS
INCLUDING COLLAGEN, FIBRIN, HEPARIN, DNA, AND ACTIN. FIBRONECTINS
ARE INVOLVED IN CELL ADHESION, CELL MOTILITY, OPSONIZATION, WOUND
HEALING, AND MAINTENANCE OF CELL SHAPE.
CC -!- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED
VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS;
TO A LESSER EXTENT HOMODIMERS.
CC -!- TISSUE SPECIFICITY: PLASMA FN (SOLUBLE DIMERIC FORM) IS SECRETED
BY HEPATOCYTES. CELLULAR FN (DIMERIC OR CROSS-LINKED MULTIMERIC
FORMS), MADE BY FIBROBLASTS, EPITHELIAL AND OTHER CELL TYPES, IS
DEPOSITED AS FIBRILS IN THE EXTRACELLULAR MATRIX.
CC -!- ALTERNATIVE PRODUCTS: EACH OF THE "EXTRA DOMAIN" & THE CONNECTING
STRAND 3 ARE PRESENT IN SOME FORMS OF FIBRONECTIN AND ABSENT IN
OTHERS. THESE DIFFERENCES ARE DUE TO ALTERNATIVE SPLICING.
CC -!- SIMILARITY: CONTAINS 12 FIBRONECTIN TYPE I DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE II DOMAINS.
CC -!- SIMILARITY: CONTAINS 16 FIBRONECTIN TYPE III DOMAINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----

```

```

DR EMBL; K00800; E18191; ALT_SEQ.
DR PIR; A26452; FNBO.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS01253; FIBRONECTIN_1; 12.
DR PFAM; PF00039; fn1; 12.
DR PFAM; PF00040; fn2; 2.
DR PFAM; PF00041; fn3; 15.
DR HSSP; P02751; 2FN2.
KW GLYCOPROTEIN; PLASMA; HEPARIN-BINDING; ACUTE PHASE; PHOSPHORYLATION;
KW CELL ADHESION; REPEAT; ALTERNATIVE SPLICING.
FT MOD_RES 1 1
FT DOMAIN 21 241
FT DOMAIN 277 577
FT DNA_BIND 876 1141
FT DOMAIN 1236 1509
FT DOMAIN 1600 1870
FT DOMAIN 1991 2216
FT DOMAIN 19 59
FT DOMAIN 64 107
FT DOMAIN 108 151
FT DOMAIN 153 197
FT DOMAIN 198 242
FT DOMAIN 275 314
FT DOMAIN 314 373
FT DOMAIN 374 438
FT DOMAIN 437 480
FT DOMAIN 485 527
FT DOMAIN 528 571
FT DOMAIN 578 669
FT DOMAIN 688 778
FT DOMAIN 779 874
FT DOMAIN 875 964
FT DOMAIN 1054 1054
FT DOMAIN 1055 1141
FT DOMAIN 1142 1234
FT DOMAIN 1235 1325
FT DOMAIN 1326 1415
FT DOMAIN 1416 1509
FT DOMAIN 1510 1599
FT DOMAIN 1600 1691
FT DOMAIN 1692 1780
FT DOMAIN 1781 1870
FT DOMAIN 1871 1990
FT DOMAIN >1870 <1982
FT DOMAIN 1982 2061
FT DOMAIN 2063 2127
FT DOMAIN 2128 2170
FT DOMAIN 2172 2215
FT SITE 1493 1495
FT DISULFID 21 47
FT DISULFID 45 56
FT DISULFID 66 94
FT DISULFID 92 104
FT DISULFID 110 138
FT DISULFID 136 148
FT DISULFID 155 184
FT DISULFID 182 194
FT DISULFID 200 229
FT DISULFID 227 239
FT DISULFID 277 304
FT DISULFID 302 311
FT DISULFID 329 355
FT DISULFID 343 370
FT DISULFID 389 415
FT DISULFID 403 430
FT DISULFID 439 467
FT DISULFID 465 477
FT DISULFID 487 514
FT DISULFID 512 524
FT DISULFID 530 558
FT DISULFID 556 568
FT DISULFID 2085 2114

```

FT DISULFID 2112 2124  
 FT DISULFID 2130 2157  
 FT DISULFID 2155 2167  
 FT DISULFID 2174 2200  
 FT DISULFID 2198 2209  
 FT DISULFID 2246 2246  
 FT DISULFID 2250 2250  
 FT CARBOHYD 339 399  
 FT CARBOHYD 497 497  
 FT CARBOHYD 511 511  
 FT CARBOHYD 846 846  
 FT CARBOHYD 976 976  
 FT CARBOHYD 1213 1213  
 FT CARBOHYD 1987 1987  
 FT CARBOHYD 1943 1943  
 FT CARBOHYD 1944 1944  
 FT MOD\_RES 2263 2263  
 SQ SEQUENCE 2265 AA; 249557 MW; B5176597 CRC32;

Query Match 97.6%; Score 40; DB 1; Length 2265;  
 Best Local Similarity 83.3%; Pred. No. 1.94e+00;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1492 GRGDSF 1497  
 QY 1 GRGESP 6  
 |||:|

RESULT 4  
 ID FINC\_HUMAN STANDARD; PRT: 2386 AA.  
 AC P02751; Q14326;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE FIBRONECTIN PRECURSOR (FN).  
 GN FN1 OR FN.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE OF 1-38 FROM N.A.  
 RX GUTMAN A., YAMADA K.M., KORNBLIHT A.;  
 RT "Human fibronectin is synthesized as a pre-propolypeptide.";  
 RL FEBS LETT. 207:145-148(1986).  
 RN [2]  
 RP SEQUENCE OF 1-49 FROM N.A.  
 RX MEDLINE; 87175578.  
 RA DEAN D.C., BOWLUS C.L., BOURGEOIS S.;  
 RT "Cloning and analysis of the promoter region of the human fibronectin gene.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 84:1876-1880(1987).  
 RN [3]  
 RP SEQUENCE OF 32-2081 AND 2113-2386 FROM N.A.  
 RX MEDLINE; 85284965.  
 RA KORNBLIHT A.R., UMZAWA K., VIBE-PEDERSEN K., BARALLE F.E.;  
 RT "Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene.";  
 RL EMBO J. 4:1755-1759(1985).  
 RN [4]  
 RP SEQUENCE OF 973-2081 AND 2113-2386 FROM N.A.  
 RX MEDLINE; 84272258.  
 RA KORNBLIHT A.R., VIBE-PEDERSEN K., BARALLE F.E.;  
 RT "Human fibronectin: cell specific alternative mRNA splicing generates polypeptide chains differing in the number of internal repeats.";  
 RL NUCLEIC ACIDS RES. 12:5853-5868(1984).  
 RN [5]  
 RP SEQUENCE OF 1594-2386 FROM N.A.  
 RX MEDLINE; 85280409.  
 RA BERNARD M.P., KOLBE M., WEIL D., CHU M.-L.;  
 RT "Human cellular fibronectin: comparison of the carboxyl-terminal portion with rat identifies primary structural domains separated by hypervariable regions.";

RL BIOCHEMISTRY 24:2698-2704(1985).  
 RN [6]  
 RP SEQUENCE OF 32-290.  
 RX MEDLINE; 84032463.  
 RA GARCIA-PARDO A., PEARLSTEIN E., FRANGIONE B.;  
 RT "Primary structure of human plasma fibronectin. The 29,000-dalton NH2-terminal domain.";  
 RL J. BIOL. CHEM. 258:12670-12674(1983).  
 RN [7]  
 RP SEQUENCE OF 309-608, AND COLLAGEN-BINDING.  
 RX MEDLINE; 87080265.  
 RA OWENS R.J., BARALLE F.E.;  
 RT "Mapping the collagen-binding site of human fibronectin by expression in Escherichia coli.";  
 RL EMBO J. 5:2825-2830(1986).  
 RN [8]  
 RP SEQUENCE OF 1441-1548.  
 RX MEDLINE; 82265604.  
 RA PIERSCHBACHER M.D., RUOSLAHTI E., SUNDELIN J., LIND P., PETERSON P.A.;  
 RT "The cell attachment domain of fibronectin. Determination of the primary structure.";  
 RL J. BIOL. CHEM. 257:9593-9597(1982).  
 RN [9]  
 RP SEQUENCE OF 1434-1537 FROM N.A.  
 RX MEDLINE; 83290929.  
 RA OLDBERG A., LINNEY E., RUOSLAHTI E.;  
 RT "Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell attachment domain in human fibronectin.";  
 RL J. BIOL. CHEM. 258:10193-10196(1983).  
 RN [10]  
 RP SEQUENCE OF 1448-1540 FROM N.A.  
 RX MEDLINE; 86111901.  
 RA OLDBERG A., RUOSLAHTI E.;  
 RT "Evolution of the fibronectin gene. Exon structure of cell attachment domain.";  
 RL J. BIOL. CHEM. 261:2113-2116(1986).  
 RN [11]  
 RP SEQUENCE OF 1712-1739 FROM N.A.  
 RX MEDLINE; 87026578.  
 RA SEKIGUCHI K., KLOS A.M., KURACHI K., YOSHITAKE S., HAKOMORI S.;  
 RT "Human liver fibronectin complementary DNAs: Identification of two different messenger RNAs possibly encoding the alpha and beta subunits of plasma fibronectin.";  
 RL BIOCHEMISTRY 25:4936-4941(1986).  
 RN [12]  
 RP CHARACTERIZATION OF FIBRIN-BINDING SITE 1.  
 RX MEDLINE; 95081153.  
 RA ROSTAGNO A., WILLIAMS M.J., BARON M., CAMPBELL I.D., GOLD L.I.;  
 RT "Further characterization of the NH2-terminal fibrin-binding site on fibronectin.";  
 RL J. BIOL. CHEM. 269:31938-31945(1994).  
 RN [13]  
 RP STRUCTURE BY NMR OF 32-92.  
 RX MEDLINE; 96069779.  
 RA POTTS J.R., PHAN I., WILLIAMS M.J., CAMPBELL I.D.;  
 RT "High-resolution structural studies of the factor XIIIa crosslinking site and the first type 1 module of fibronectin.";  
 RL NAT. STRUCT. BIOL. 2:946-950(1995).  
 RN [14]  
 RP STRUCTURE BY NMR OF 182-275.  
 RX MEDLINE; 94141923.  
 RA WILLIAMS M.J., PHAN I., HARVEY T.S., ROSTAGNO A., GOLD L.I., CAMPBELL I.D.;  
 RT "Solution structure of a pair of fibronectin type 1 modules with fibrin binding activity.";  
 RL J. MOL. BIOL. 235:1302-1311(1994).  
 RN [15]  
 RP STRUCTURE BY NMR OF 406-464.  
 RX MEDLINE; 98179558.  
 RA STICHT H., PICKFORD A.R., POTTS J.R., CAMPBELL I.D.;  
 RT "Solution structure of the glycosylated second type 2 module of fibronectin.";  
 RL J. MOL. BIOL. 276:177-187(1998).

[16] RP STRUCTURE BY NMR OF 1447-1540.  
 RX MEDLINE; 93046665.  
 RA MAIN A.L., HARVEY T.S., BARON M., BOYD J., CAMPBELL I.D.;  
 RT "The three-dimensional structure of the tenth type III module of  
 FT fibronectin: an insight into RGD-mediated interactions.";  
 RL CELL 71:671-678(1992).  
 RN [17]  
 RP STRUCTURE BY NMR OF 1447-1540.  
 RX MEDLINE; 92162710.  
 RA BARON M., MAIN A.L., DRISCOLL P.C., MARDON H.J., BOYD J.,  
 RA CAMPBELL I.D.;  
 RT "1H NMR assignment and secondary structure of the cell adhesion type  
 FT III module of fibronectin.";  
 RL BIOCHEMISTRY 31:2068-2073(1992).  
 RN [18]  
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 1447-1535.  
 RX MEDLINE; 94166075.  
 RA DICKINSON C.D., VEERAPANDIAN B., DAI X.-P., HAMLIN R.C., XUONG N.-H.,  
 RA RUOSLAHTI E., ELY K.R.;  
 RT "Crystal structure of the tenth type III cell adhesion module of  
 FT human fibronectin.";  
 RL J. MOL. BIOL. 236:1079-1092(1994).  
 CC -!- FUNCTION: FIBRONECTINS BIND CELL SURFACES AND VARIOUS COMPOUNDS  
 CC INCLUDING COLLAGEN, FIBRIN, HEPARIN, DNA, AND ACTIN. FIBRONECTINS  
 CC ARE INVOLVED IN CELL ADHESION, CELL MOTILITY, OPSONIZATION, WOUND  
 CC HEALING, AND MAINTENANCE OF CELL SHAPE.  
 CC -!- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED  
 CC VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS;  
 CC TO A LESSER EXTENT HOMODIMERS.  
 CC -!- TISSUE SPECIFICITY: PLASMA FN (SOLUBLE DIMERIC FORM) IS SECRETED  
 CC BY HEPATOCYTES. CELLULAR FN (DIMERIC OR CROSS-LINKED MULTIMERIC  
 CC FORMS), MADE BY FIBROBLASTS, EPITHELIAL AND OTHER CELL TYPES, IS  
 CC DEPOSITED AS FIBRILS IN THE EXTRACELLULAR MATRIX.  
 CC -!- ALTERNATIVE PRODUCTS: EACH OF THE "EXTRA DOMAIN" & THE CONNECTING  
 CC STRAND 3 ARE PRESENT IN SOME FORMS OF FIBRONECTIN AND ABSENT IN  
 CC OTHERS. THESE DIFFERENCES ARE DUE TO ALTERNATIVE SPLICING.  
 CC -!- SIMILARITY: CONTAINS 12 FIBRONECTIN TYPE I DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE II DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 17 FIBRONECTIN TYPE III DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M15801; G553293; -.  
 DR EMBL; X02761; G31397; -.  
 DR EMBL; K00055; G182683; -.  
 DR EMBL; M10905; G182697; -.  
 DR EMBL; M12549; G553294; -.  
 DR EMBL; M14059; G182703; -.  
 DR PIR; A26460; FNHU.  
 DR PDB; ITTF; 31-JAN-94.  
 DR PDB; ITTG; 31-JAN-94.  
 DR PDB; IFNA; 30-APR-94.  
 DR PDB; IFNF; 29-JAN-96.  
 DR PDB; IFBR; 15-OCT-95.  
 DR PDB; 2FN2; PRELIMINARY.  
 DR MIM; I35600; -.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS00023; FIBRONECTIN\_2; 2.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 12.  
 DR PFAM; PF00039; fn1; 12.  
 DR PFAM; PF00040; fn2; 2.  
 DR PFAM; PF00041; fn3; 16.  
 KW GLYCOPROTEIN; PLASMA; HEPARIN-BINDING; ACUTE PHASE; PHOSPHORYLATION;  
 KW SULFATATION; CELL ADHESION; REPEAT; ALTERNATIVE SPLICING; SIGNAL;  
 FT 3D-STRUCTURE. 1 31

FT CHAIN 32 2386 FIBRONECTIN.  
 FT DOMAIN 52 272 FIBRIN- AND HEPARIN-BINDING 1.  
 FT DOMAIN 308 608 COLLAGEN-BINDING.  
 FT DNA\_BIND 907 1172 CELL-ATTACHMENT.  
 FT DOMAIN 1267 1540 HEPARIN-BINDING 2.  
 FT DOMAIN 1721 1991 FIBRIN-BINDING 2.  
 FT DOMAIN 2206 2337 CELL ATTACHMENT SITE.  
 FT SITE 1524 1526 FIBRONECTIN TYPE-I 1.  
 FT DOMAIN 50 90 FIBRONECTIN TYPE-I 2.  
 FT DOMAIN 95 138 FIBRONECTIN TYPE-I 2.  
 ...  
 Note: remainder of annotations omitted.  
 Query Match 97.6%; Score 40; DB 1; Length 2386;  
 Best Local Similarity 83.3%; Pred. No. 1.94e+00;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Db 1523 GRGSP 1528  
 Qy 1 GRGSP 6  
 |||:|  
 RESULT 5  
 ID FUNC\_MOUSE STANDARD; PRT: 2477 AA.  
 AC P11276; Q61569; Q61567; Q64233;  
 DT 01-JUL-1989 (REL. 11, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE FIBRONECTIN PRECURSOR (FN) (FRAGMENTS).  
 GN FNI.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE OF 1-28 FROM N.A.  
 RC TISSUE-LIVER;  
 RX MEDLINE; 94131313.  
 RA POLLY P., NICHOLSON R.C.;  
 RT "Sequence of the mouse fibronectin-encoding gene promoter region.";  
 RL GENE 137:353-354(1993).  
 RN [2]  
 RP SEQUENCE OF 562-834 FROM N.A.  
 RC STRAIN=NMRL;  
 RX MEDLINE; 95403556.  
 RA TALTS J.F., WELLER A., TIMPL R., EKBLOM M., EKBLOM P.;  
 RT "Regulation of mesenchymal extracellular matrix protein synthesis by  
 FT transforming growth factor-beta and glucocorticoids in tumor  
 FT stroma.";  
 RL J. CELL SCI. 108:2153-2162(1995).  
 RN [3]  
 RP SEQUENCE OF 899-2376 FROM N.A.  
 RA GORSKI G., AROS M., NORTON P.;  
 RL SUBMITTED (DEC-1995) TO EMBL/GENEBANK/DBBJ DATA BANKS.  
 RN [4]  
 RP SEQUENCE OF 2375-2477 FROM N.A.  
 RX MEDLINE; 88124987.  
 RA BLATTI S.P., FOSTER D.N., RANGANTHAN G., MOSES H.L., GETZ M.J.;  
 RT "Induction of fibronectin gene transcription and mRNA is a primary  
 RT response to growth-factor stimulation of ABR-28 cells.";  
 RL PROC. NACL. ACAD. SCI. U.S.A. 85:1119-1123(1988).  
 RN [5]  
 RP SEQUENCE OF 2375-2477 FROM N.A.  
 RC TISSUE-KIDNEY;  
 RX MEDLINE; 93011702.  
 RA KHANDJIAN E.W., SALOMON C., LEONARD N., TREMBLAY S., TURLER H.;  
 RT "Fibronectin gene expression in proliferating, quiescent, and SV40-  
 RT infected mouse kidney cells.";  
 RL EXP. CELL RES. 202:464-470(1992).  
 RN [6]  
 RP STRUCTURE BY NMR OF 1447-1630.  
 RX MEDLINE; 98202578.  
 RA COPIE V., TOMITA-Y., AKIYAMA S.K., AOTA S., YAMADA K.M., VENABLE R.M.,  
 RA PASTOR R.W., KRUEGER S., TORCHIA D.A.;

"Solution structure and dynamics of linked cell attachment modules of mouse fibronectin containing the RGD and synergy regions: comparison with the human fibronectin crystal structure."

J. MOL. BIOL. 277:663-682(1998).

!- FUNCTION: FIBRONECTIN BIND CELL SURFACES AND VARIOUS COMPOUNDS INCLUDING COLLAGEN, FIBRIN, HEPARIN, DNA, AND ACTIN. FIBRONECTINS ARE INVOLVED IN CELL ADHESION, CELL MOTILITY, OPSONIZATION, WOUND HEALING, AND MAINTENANCE OF CELL SHAPE.

!- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS; TO A LESSER EXTENT HOMODIMERS.

!- TISSUE SPECIFICITY: PLASMA FN (SOLUBLE DIMERIC FORM) IS SECRETED BY HEPATOCYTES. CELLULAR FN (DIMERIC OR CROSS-LINKED MULTIMERIC FORMS), MADE BY FIBROBLASTS, EPITHELIAL AND OTHER CELL TYPES, IS DEPOSITED AS FIBRILS IN THE EXTRACELLULAR MATRIX.

!- ALTERNATIVE PRODUCTS: EACH OF THE "EXTRA DOMAIN" & THE CONNECTING STRAND 3 ARE PRESENT IN SOME FORMS OF FIBRONECTIN AND ABSENT IN OTHERS. THESE DIFFERENCES ARE DUE TO ALTERNATIVE SPLICING.

!- SIMILARITY: CONTAINS 12 FIBRONECTIN TYPE I DOMAINS.

!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE II DOMAINS.

!- SIMILARITY: CONTAINS 18 FIBRONECTIN TYPE III DOMAINS.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; 222729; G297912; -  
 EMBL; X82402; E122892; -  
 DR EMBL; X93167; E210659; -  
 DR EMBL; M18194; G387159; -  
 DR EMBL; S45680; G256716; -  
 DR PIR; A31371; A31371;  
 DR PIR; C60597; C60597;  
 DR PDB; 2FN; 29-APR-98.  
 DR PDB; 2MFN; 28-APR-98.  
 DR MGD; MGI:95566; FN1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS00023; FIBRONECTIN\_2; PARTIAL.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 3.  
 DR PFAM; PF00039; fn1; 4.  
 DR PFAM; PF00041; fn3; 16.  
 KW GLYCOPROTEIN; PLASMA; HEPARIN-BINDING; ACUTE PHASE; PHOSPHORYLATION;  
 KW SULFATATION; CELL ADHESION; REPEAT; ALTERNATIVE SPLICING; SIGNAL;  
 3D-STRUCTURE.  
 FT SIGNAL 1 32 BY SIMILARITY.  
 FT CHAIN 33 2477 FIBRONECTIN.  
 FT DOMAIN 53 273 FIBRIN- AND HEPARIN-BINDING 1.  
 FT DOMAIN 308 608 COLLAGEN-BINDING.  
 FT DNA\_BIND 906 1171  
 FT DOMAIN 1357 1630  
 FT DOMAIN 1811 2081  
 FT DOMAIN 2296 2427  
 FT DOMAIN 51 96  
 FT DOMAIN 96 140 FIBRONECTIN TYPE-I 1.  
 FT DOMAIN 140 185 FIBRONECTIN TYPE-I 2.  
 FT DOMAIN 185 230 FIBRONECTIN TYPE-I 3.  
 FT DOMAIN 230 272 FIBRONECTIN TYPE-I 4.  
 FT DOMAIN 306 343 FIBRONECTIN TYPE-I 5.  
 FT DOMAIN 345 404 FIBRONECTIN TYPE-II 1.  
 FT DOMAIN 405 469 FIBRONECTIN TYPE-II 2.  
 FT DOMAIN 468 516 FIBRONECTIN TYPE-I 7.  
 FT DOMAIN 516 559 FIBRONECTIN TYPE-I 8.  
 FT DOMAIN 559 602 FIBRONECTIN TYPE-I 9.  
 FT DOMAIN 609 706 FIBRONECTIN TYPE-III 1.  
 FT DOMAIN 707 808 FIBRONECTIN TYPE-III 2.  
 FT DOMAIN 809 903 FIBRONECTIN TYPE-III 3.  
 FT DOMAIN 994 1084 FIBRONECTIN TYPE-III 4.  
 FT DOMAIN 995 1084 FIBRONECTIN TYPE-III 5.  
 FT DOMAIN 1085 1172 FIBRONECTIN TYPE-III 6.

FT DOMAIN 1173 1264 FIBRONECTIN TYPE-III 7.  
 FT DOMAIN 1265 1355 FIBRONECTIN TYPE-III 8 (EXTRA DOMAIN 1).  
 FT DOMAIN 1356 1446 FIBRONECTIN TYPE-III 9.  
 FT DOMAIN 1447 1536 FIBRONECTIN TYPE-III 10.  
 FT DOMAIN 1537 1630 FIBRONECTIN TYPE-III 11.  
 FT DOMAIN 1631 1720 FIBRONECTIN TYPE-III 12.  
 FT DOMAIN 1721 1810 FIBRONECTIN TYPE-III 13 (EXTRA DOMAIN 2).  
 FT DOMAIN 1811 1902 FIBRONECTIN TYPE-III 14.  
 FT DOMAIN 1903 1991 FIBRONECTIN TYPE-III 15.  
 FT DOMAIN 1992 2081 FIBRONECTIN TYPE-III 16.  
 FT DOMAIN 2082 2201 CONNECTING STRAND 3 (CS-3) (V REGION).  
 FT DOMAIN >2202 FIBRONECTIN TYPE-III 17.  
 FT DOMAIN 2202 2283 FIBRONECTIN TYPE-III 18.  
 FT DOMAIN 2294 2338 FIBRONECTIN TYPE-I 10.  
 FT DOMAIN 2339 2381 FIBRONECTIN TYPE-I 11.  
 FT DOMAIN 2383 2426 FIBRONECTIN TYPE-I 12.  
 FT SITE 1614 1616 CELL ATTACHMENT SITE.  
 FT SITE 2181 2183 CELL ATTACHMENT SITE.  
 FT DISULFID 561 589 BY SIMILARITY.  
 FT DISULFID 587 599 BY SIMILARITY.  
 FT DISULFID 2296 2325 BY SIMILARITY.  
 FT DISULFID 2323 2335 BY SIMILARITY.  
 FT DISULFID 2341 2368 BY SIMILARITY.  
 FT DISULFID 2366 2378 BY SIMILARITY.  
 FT DISULFID 2385 2409 BY SIMILARITY.  
 FT DISULFID 2407 2423 BY SIMILARITY.  
 FT DISULFID 2458 2458 INTERCHAIN (WITH 2462 OF OTHER CHAIN).  
 FT DISULFID 2462 2462 INTERCHAIN (WITH 2458 OF OTHER CHAIN).  
 FT CARBOHYD 1006 1006 POTENTIAL.  
 FT CARBOHYD 1243 1243 POTENTIAL.  
 FT CARBOHYD 1290 1290 POTENTIAL.  
 FT CARBOHYD 2198 2198 POTENTIAL.  
 FT MOD\_RES 2475 2475 PHOSPHORYLATION (BY SIMILARITY).  
 FT CONFLICT 2440 2440 N -> T (IN REF 5).  
 SQ SEQUENCE 2477 AA; 271416 MW; 744ECC2B CRC32;  
 Query Match 97.6%; Score 40; DB 1; Length 2477;  
 Best Local Similarity 83.3%; Pred. No. 1.94e+00;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1613 GRGDSF 1618  
 |||:|  
 Qy 1 GRGESP 6

RESULT 6  
 ID FUNC\_RAT STANDARD; PRT; 2477 AA.  
 AC P04937;  
 DT 13-AUG-1987 (REL. 05, CREATED)  
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE FIBRONECTIN PRECURSOR (FN).  
 GN FN1.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FISHER; TISSUE-LIVER;  
 RX MEDLINE; 88054951.  
 RA SCHWARZBAUER J.E., PATEL R.S., FONDA D., HYNES R.O.;  
 RT "Multiple sites of alternative splicing of the rat fibronectin gene transcript."  
 RL EMBO J. 6:2573-2580(1987).  
 [2]  
 RP SEQUENCE OF 1-139 AND 2382-2477 FROM N.A.  
 RC STRAIN-FISHER; TISSUE-LIVER;  
 RX MEDLINE; 88054950.  
 RA PATEL R.S., ODERMATT E., SCHWARZBAUER J.E., HYNES R.O.;  
 RT "Organization of the fibronectin gene provides evidence for exon shuffling during evolution."  
 RL EMBO J. 6:2565-2572(1987).  
 [3]

RP SEQUENCE OF 1586-2477 FROM N.A.  
 RX MEDLINE; 84082067.  
 RA SCHWARZBAUER J.E., TAMKUN J.W., LEWISCHKA I.R., HYNES R.O.;  
 RT "Three different fibronectin mRNAs arise by alternative splicing  
 within the coding region.";  
 RL CELL 35;421-431(1983).  
 CC !- FUNCTION: FIBRONECTINS BIND CELL SURFACES AND VARIOUS COMPOUNDS  
 INCLUDING COLLAGEN, FIBRIN, HEPARIN, DNA, AND ACTIN. FIBRONECTINS  
 ARE INVOLVED IN CELL ADHESION, CELL MOTILITY, OPSONIZATION, WOUND  
 HEALING, AND MAINTENANCE OF CELL SHAPE.  
 CC !- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED  
 VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS;  
 TO A LESSER EXTENT HOMODIMERS.  
 CC !- TISSUE SPECIFICITY: PLASMA FN (SOLUBLE DIMERIC FORM) IS SECRETED  
 BY HEPATOCYTES. CELLULAR FN (DIMERIC OR CROSS-LINKED MULTIMERIC  
 FORMS), MADE BY FIBROBLASTS, EPITHELIAL AND OTHER CELL TYPES, IS  
 CC DEPOSITED AS FIBRILS IN THE EXTRACELLULAR MATRIX.  
 CC !- ALTERNATIVE PRODUCTS: EACH OF THE "EXTRA DOMAIN" & THE CONNECTING  
 STRAND 3 ARE PRESENT IN SOME FORMS OF FIBRONECTIN AND ABSENT IN  
 CC OTHERS. THESE DIFFERENCES ARE DUE TO ALTERNATIVE SPLICING.  
 CC !- SIMILARITY: CONTAINS 12 FIBRONECTIN TYPE I DOMAINS.  
 CC !- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE II DOMAINS.  
 CC !- SIMILARITY: CONTAINS 18 FIBRONECTIN TYPE III DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X15906; G56164; -  
 DR EMBL; L29191; G204156; -  
 DR EMBL; L00191; G204156; JOINED.  
 DR EMBL; L29191; G204157; -  
 DR EMBL; L00191; G204157; JOINED.  
 DR EMBL; L29191; G204158; -  
 DR EMBL; L00191; G204158; JOINED.  
 DR EMBL; X05831; G56156; -  
 DR EMBL; X05832; G769820; -  
 DR EMBL; X05833; G773260; -  
 DR EMBL; X05834; G773261; -  
 DR PIR; A27252; A27252.  
 DR PIR; S00459; S00459.  
 DR PIR; S14428; S14428.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS00023; FIBRONECTIN\_2; 2.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 12.  
 DR PFAM; PF00039; fn1; 12.  
 DR PFAM; PF00040; fn2; 2.  
 DR PFAM; PF00041; fn3; 17.  
 DR HSP; P02751; 1FBR.  
 KW GLYCOPROTEIN; PLASMA; HEPARIN-BINDING; ACUTE PHASE; PHOSPHORYLATION;  
 KW SULFATATION; CELL ADHESION; REPEAT; ALTERNATIVE SPLICING; SIGNAL.  
 FT SIGNAL  
 FT CHAIN 1 32  
 FT CHAIN 33 2477 FIBRONECTIN.  
 FT DOMAIN 53 273 FIBRIN- AND HEPARIN-BINDING 1.  
 FT DOMAIN 308 608 COLLAGEN-BINDING.  
 FT DNA\_BIND 906 1171  
 FT DOMAIN 1337 1630 CELL-ATTACHMENT.  
 FT DOMAIN 1811 2081 HEPARIN-BINDING 2.  
 FT DOMAIN 2296 2427 FIBRIN-BINDING 2.  
 FT DOMAIN 51 91 FIBRONECTIN TYPE-I 1.  
 FT DOMAIN 96 139 FIBRONECTIN TYPE-I 2.  
 FT DOMAIN 140 183 FIBRONECTIN TYPE-I 3.  
 FT DOMAIN 185 229 FIBRONECTIN TYPE-I 4.  
 FT DOMAIN 230 274 FIBRONECTIN TYPE-I 5.  
 FT DOMAIN 306 345 FIBRONECTIN TYPE-I 6.  
 FT DOMAIN 345 404 FIBRONECTIN TYPE-II 1.  
 FT DOMAIN 405 469 FIBRONECTIN TYPE-II 2.  
 FT DOMAIN 468 511 FIBRONECTIN TYPE-I 7.  
 FT DOMAIN 516 FIBRONECTIN TYPE-I 8.

FT DOMAIN 559 706 FIBRONECTIN TYPE-I 9.  
 FT DOMAIN 707 808 FIBRONECTIN TYPE-III 1.  
 FT DOMAIN 809 903 FIBRONECTIN TYPE-III 2.  
 FT DOMAIN 904 994 FIBRONECTIN TYPE-III 3.  
 FT DOMAIN 995 1084 FIBRONECTIN TYPE-III 4.  
 FT DOMAIN 1085 1172 FIBRONECTIN TYPE-III 5.  
 FT DOMAIN 1173 1264 FIBRONECTIN TYPE-III 6.  
 FT DOMAIN 1265 1355 FIBRONECTIN TYPE-III 7.  
 FT DOMAIN 1356 1446 FIBRONECTIN TYPE-III 8 (EXTRA DOMAIN 1).  
 FT DOMAIN 1447 1536 FIBRONECTIN TYPE-III 9.  
 FT DOMAIN 1537 1630 FIBRONECTIN TYPE-III 10.  
 FT DOMAIN 1631 1720 FIBRONECTIN TYPE-III 11.  
 FT DOMAIN 1721 1810 FIBRONECTIN TYPE-III 12.  
 FT DOMAIN 1811 1902 FIBRONECTIN TYPE-III 13 (EXTRA DOMAIN 2).  
 FT DOMAIN 1903 1991 FIBRONECTIN TYPE-III 14.  
 FT DOMAIN 1992 2081 FIBRONECTIN TYPE-III 15.  
 FT DOMAIN 2082 2201 FIBRONECTIN TYPE-III 16.  
 FT DOMAIN >2081 <2202 CONNECTING STRAND 3 (CS-3) (V REGION).  
 FT DOMAIN 2202 2283 FIBRONECTIN TYPE-III 17.  
 FT DOMAIN 2283 2338 FIBRONECTIN TYPE-III 18.  
 FT DOMAIN 2338 2381 FIBRONECTIN TYPE-I 10.  
 FT DOMAIN 2381 2426 FIBRONECTIN TYPE-I 11.  
 FT DOMAIN 2426 2462 FIBRONECTIN TYPE-I 12.  
 FT SITE 1614 1616 CELL ATTACHMENT SITE.  
 FT SITE 2381 2383 CELL ATTACHMENT SITE.  
 FT DISULFID 53 79 BY SIMILARITY.  
 FT DISULFID 77 88 BY SIMILARITY.  
 FT DISULFID 98 126 BY SIMILARITY.  
 FT DISULFID 124 136 BY SIMILARITY.  
 FT DISULFID 142 170 BY SIMILARITY.  
 FT DISULFID 168 180 BY SIMILARITY.  
 FT DISULFID 187 216 BY SIMILARITY.  
 FT DISULFID 214 226 BY SIMILARITY.  
 FT DISULFID 232 261 BY SIMILARITY.  
 FT DISULFID 259 271 BY SIMILARITY.  
 FT DISULFID 308 335 BY SIMILARITY.  
 FT DISULFID 333 342 BY SIMILARITY.  
 FT DISULFID 360 386 BY SIMILARITY.  
 FT DISULFID 374 401 BY SIMILARITY.  
 FT DISULFID 420 446 BY SIMILARITY.  
 FT DISULFID 434 461 BY SIMILARITY.  
 FT DISULFID 470 498 BY SIMILARITY.  
 FT DISULFID 496 508 BY SIMILARITY.  
 FT DISULFID 518 545 BY SIMILARITY.  
 FT DISULFID 543 555 BY SIMILARITY.  
 FT DISULFID 561 589 BY SIMILARITY.  
 FT DISULFID 587 599 BY SIMILARITY.  
 FT DISULFID 2296 2325 BY SIMILARITY.  
 FT DISULFID 2323 2335 BY SIMILARITY.  
 FT DISULFID 2341 2368 BY SIMILARITY.  
 FT DISULFID 2366 2378 BY SIMILARITY.  
 FT DISULFID 2385 2409 BY SIMILARITY.  
 FT DISULFID 2407 2423 BY SIMILARITY.  
 FT DISULFID 2458 2458 INTERCHAIN (WITH 2462 OF OTHER CHAIN).  
 FT DISULFID 2462 2462 INTERCHAIN (WITH 2458 OF OTHER CHAIN).  
 FT CARBOHYD 430 430 POTENTIAL.  
 FT CARBOHYD 528 528 POTENTIAL.  
 FT CARBOHYD 542 542 POTENTIAL.  
 FT CARBOHYD 876 876 POTENTIAL.  
 FT CARBOHYD 1006 1006 POTENTIAL.  
 FT CARBOHYD 1243 1243 POTENTIAL.  
 FT CARBOHYD 1290 1290 POTENTIAL.  
 FT CARBOHYD 2198 2198 POTENTIAL.  
 FT MOD\_RES 2475 2475 PHOSPHORYLATION (BY SIMILARITY).  
 FT VARSPLIC 1720 1809 MISSING (IN ALTERNATIVE SPLICED FORM).  
 FT VARSPLIC 2082 2106 MISSING (IN CLONE LAMBDA-RLF4-5).  
 FT VARSPLIC 2082 2200 MISSING (IN CLONE LAMBDA-RLF6).  
 FT CONFLICT 2318 2318 G -> A (IN REF. 3).  
 SQ SEQUENCE 2477 AA; 272510 MW; 093A8F76 CRC32;

Query Match 97.6%; Score 40; DB 1; Length 2477;  
 Best Local Similarity 83.3%; Pred. No. 1.94e+00;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



Db 1614 GRGDSP 1619



```

QY 1 GRGESP 6
|||||
RESULT 8
ID YIK3_YEAST STANDARD; PRT; 425 AA.
AC P04087;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHEICAL 48.3 KD PROTEIN IN MOB1-SGAL INTERGENIC REGION.
GN YIL103W.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-S288C / AB972;
RC BARRELL B.G., BADCOCK K., BANKIER A.T., BOWMAN S., BROWN D.,
RA CHURCHER C.M., CONNOR R., CORSEY T., DEAR S., DEVLIN K., FRASER A.,
RA GENTLES S., HAMLYN N., HORSNELL T.S., HUNT S., JAGELS K., JONES M.,
RA LOUIS E., LYE G., MOULE T., ODELL C., PEARSON D.,
RA RAJANDREAM M.A., RILES L., ROWLEY N., SKELTON J., SMITH V.,
RA WALSH S.V., WHITEHEAD S.;
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- SIMILARITY: TO YEAST DIPHTHERIA TOXIN RESISTANCE PROTEIN 2 (DPH2),
CC TO S.POMBE SPAC13F4.15C, TO C.ELEGANS C09G5.2, C14B1.5 AND
CC M.JANNASCHII MJ0483.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z38125; G558703; -.
CC EMBL; Z47047; G763243; -.
CC PIR; S48469; S48469.
CC HYPOTHEICAL PROTEIN.
CC SEQUENCE 425 AA; 48310 MW; 28921DCE CRC32;
Query Match 95.1%; Score 39; DB 1; Length 425;
Best Local Similarity 83.3%; Pred. No. 3.67e+00;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 415 GRGETP 420
|||||
QY 1 GRGESP 6
RESULT 9
ID Y863_SYNY3 STANDARD; PRT; 463 AA.
AC P73754;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE HYPOTHEICAL 50.4 KD PROTEIN SLR0863.
GN SLR0863.
OS SYNECHOCYSTIS SP. (STRAIN PCC 6803).
OC BACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCYSTIS.
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE; 97061201.
RC KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
RA HOSOUCHI T., MURAKAWA A., MURAKAWA A., NAKAZAKI N., NARUO K.,
RA OKUMURA S., SHIMPO S., TAKEUCHI C., WADA T., WATANABE A.,
RA YAMADA M., YASUDA M., TABATA S.;
RL "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
DNA RES. 3:109-136(1996).
CC -1- SIMILARITY: BELONGS TO THE TLDD/PMBA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90909; G1652888; -.
CC HYPOTHEICAL PROTEIN.
CC SEQUENCE 463 AA; 50369 MW; FE40976C CRC32;
Query Match 95.1%; Score 39; DB 1; Length 463;
Best Local Similarity 83.3%; Pred. No. 3.67e+00;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 352 GRGETP 357
|||||
QY 1 GRGESP 6
RESULT 10
ID FILA_MOUSE STANDARD; PRT; 336 AA.
AC P11088;
DT 01-JUL-1989 (REL. 11, CREATED)
DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE FILAGGRIN (FRAGMENT).
GN FLG.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE; 88058903.
RC ROTHNAGEL J.A., MEHREL T., IDLER W.W., ROOP D.R., STEINERT P.M.;
RA "The gene for mouse epidermal filaggrin precursor. Its partial
RT characterization, expression, and sequence of a repeating filaggrin
RT unit.";
RL J. BIOL. CHEM. 262:15643-15648(1987).
RN [2]
RP REVISIONS.
RC ROTHNAGEL J.A.;
RL SUBMITTED (SEP-1988) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: FILAGGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS
CC AND PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE
CC FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN
CC EPIDERMIS.
CC -1- PTM: FILAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES
CC OF 248 AA, WHICH ARE NOT SEPARATED BY "LARGE LINKER". THE
CC PRECURSOR IS DEPOSITED AS KERATOHYALIN GRANULES, DURING TERMINAL
CC DIFFERENTIATION IT IS DEPHOSPHORYLATED & PROTEOLYTICALLY CLEAVED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J03458; G387157; -.
CC PIR; A28444; A28444.
CC MGD; MGI:955553; FLG.
CC PHOSPHORYLATION; POLYPROTEIN; DEVELOPMENTAL PROTEIN.
CC NON_TER 1
CC SEQUENCE 336 AA; 35678 MW; 44E31E65 CRC32;
Query Match 92.7%; Score 38; DB 1; Length 336;

```

Best Local Similarity 83.3%; Pred. No. 6.87e+00; Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 246 GRGSP 251  
|||||  
QY 1 GRGSP 6

RESULT 11  
ID HUPK\_RHLV STANDARD; PRT; 370 AA.  
AC P28153;  
DT 01-JUL-1993 (REL. 26, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE HYDROGENASE EXPRESSION/FORMATION PROTEIN HUPK.  
GN HUPK.  
OS RHIZOBium LEGUMINOSARUM (BIOVAR VICIAE).  
OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; RHIZOBIACEAE GROUP;  
OC RHIZOBIACEAE; RHIZOBIUM.  
FN [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN-128C53;  
RX MEDLINE; 93108466.  
RA REY L., HIDALGO E., PALACIOS J.M., RUIZ-ARGUESO T.;  
RT "Nucleotide sequence and organization of an H2-uptake gene cluster  
from Rhizobium leguminosarum bv. viciae containing a rubredoxin-like  
gene and four additional open reading frames.";  
RL J. MOL. BIOL. 228:998-1002(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 95020662.  
RA IMPERIAL J., REY L., RUIZ-ARGUESO T.;  
RT "HupK, a hydrogenase-ancillary protein from Rhizobium leguminosarum,  
shares structural motifs with the large subunit of NifE hydrogenases  
and could be a scaffolding protein for hydrogenase metal cofactor  
assembly.";  
RL MOL. MICROBIOL. 9:1305-1306(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN-B10;  
RA SCHMITT H., KOKOTEK W., THUERING H., KERL V., BAUER T., FUCHS D.,  
TICHY H., LOTZ W.;  
RT "Submitted (SEP-1994) to EMBL/GENBANK/DBJ DATA BANKS.  
-1- SIMILARITY: BELONGS TO THE HUPK FAMILY.  
-2- This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-3- EMBL; X52974; G48731; -  
DR EMBL; X36981; G536808; -  
FT CONFLICT 240 240 A -> R (IN REF. 1).  
SQ SEQUENCE 370 AA; 38751 MW; 3F9E2409 CRC32;

Query Match 92.7%; Score 38; DB 1; Length 370;  
Best Local Similarity 66.7%; Pred. No. 6.87e+00; Indels 0; Gaps 0;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 163 GRGDTF 168  
|||||  
QY 1 GRGESP 6

RESULT 12  
ID CP83\_PIG STANDARD; PRT; 492 AA.  
AC Q02390;  
DT 01-JUL-1993 (REL. 26, CREATED)  
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

Query Match 92.7%; Score 38; DB 1; Length 492;  
Best Local Similarity 83.3%; Pred. No. 6.87e+00; Indels 0; Gaps 0;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 484 GRGEGP 489  
|||||  
QY 1 GRGESP 6

RESULT 13  
ID CP81\_PIG STANDARD; PRT; 492 AA.  
AC P15540;  
DT 01-APR-1990 (REL. 14, CREATED)  
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE CYTOCHROME P450 XXIA1 (EC 1.14.99.10) (STERIOD 21-HYDROXYLASE)  
DE (P450-C21).  
GN CYP21A1.  
OS SUS SCROFA (PIG).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC ARTIODACTYLA; SUIFORMES; SUINA; SUIDAE; SUS.  
RN [1]  
RP SEQUENCE.  
RX MEDLINE; 87212013.  
RA HANIU M., YANAGIBASHI K., HALL P.F., SHIVELY J.E.;  
RT "Complete amino acid sequence of 21-hydroxylase cytochrome P-450 from  
porcine adrenal microsomes.";  
RL ARCH. BIOCHEM. BIOPHYS. 254:380-384(1987).  
CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE  
MONOOXYGENASES. THEY OXIDIZE A VARIETY OF STRUCTURALLY UNRELATED  
COMPOUNDS, INCLUDING STEROIDS, FATTY ACIDS, AND XENOBIOTICS.

DE CYTOCHROME P450 XXIA3 (EC 1.14.99.10) (STERIOD 21-HYDROXYLASE)  
DE (P450-C21).  
GN CYP21A3.  
OS SUS SCROFA (PIG).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC ARTIODACTYLA; SUIFORMES; SUINA; SUIDAE; SUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 93129614.  
RA BURGHELLE-MAYEUR C., GEFEROTIN C., VAIKMAN M.;  
RT "Sequences of the swine 21-hydroxylase gene (CYP21) and a portion of  
the opposite-strand overlapping gene of unknown function previously  
described in human.";  
RL BIOCHIM. BIOPHYS. ACTA 1171:153-161(1992).  
CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE  
MONOOXYGENASES. THEY OXIDIZE A VARIETY OF STRUCTURALLY UNRELATED  
COMPOUNDS, INCLUDING STEROIDS, FATTY ACIDS, AND XENOBIOTICS.  
CC -1- CATALYTIC ACTIVITY: A STEROID + AH(2) + O(2) = A 21-HYDROXYSTEROID  
+ A + H(2)O.  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.  
CC -1- DOMAIN: THE LEUCINE-RICH HYDROPHOBIC AMINO ACID N-TERMINAL REGION  
PROBABLY HELPS TO ANCHOR THE PROTEIN TO THE MICROSMAL MEMBRANE.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC -2- This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-3- EMBL; M83939; G164560; -  
DR EMBL; M83939; G164560; -  
DR PIR; S28169; S28169  
DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
DR PFAM; PF00087; P450; 1.  
KW OXIDOREDUCTASE; MONOOXYGENASE; MEMBRANE; HEME; MICROSOME;  
KW ENDOPLASMIC RETICULUM; STEROID-GENESIS; STEROID-BINDING.  
FT BINDING 427 427 HEME (BY SIMILARITY).  
FT DOMAIN 341 357 STEROID-BINDING (POTENTIAL).  
SQ SEQUENCE 492 AA; 55619 MW; 63A1CCEFCRC32;

Query Match 92.7%; Score 38; DB 1; Length 492;  
Best Local Similarity 83.3%; Pred. No. 6.87e+00; Indels 0; Gaps 0;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 484 GRGEGP 489  
|||||  
QY 1 GRGESP 6

RESULT 13  
ID CP81\_PIG STANDARD; PRT; 492 AA.  
AC P15540;  
DT 01-APR-1990 (REL. 14, CREATED)  
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE CYTOCHROME P450 XXIA1 (EC 1.14.99.10) (STERIOD 21-HYDROXYLASE)  
DE (P450-C21).  
GN CYP21A1.  
OS SUS SCROFA (PIG).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC ARTIODACTYLA; SUIFORMES; SUINA; SUIDAE; SUS.  
RN [1]  
RP SEQUENCE.  
RX MEDLINE; 87212013.  
RA HANIU M., YANAGIBASHI K., HALL P.F., SHIVELY J.E.;  
RT "Complete amino acid sequence of 21-hydroxylase cytochrome P-450 from  
porcine adrenal microsomes.";  
RL ARCH. BIOCHEM. BIOPHYS. 254:380-384(1987).  
CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE  
MONOOXYGENASES. THEY OXIDIZE A VARIETY OF STRUCTURALLY UNRELATED  
COMPOUNDS, INCLUDING STEROIDS, FATTY ACIDS, AND XENOBIOTICS.

```

CC -!- CATALYTIC ACTIVITY: A STEROID + AH(2) + O(2) - A 21-HYDROXYSTEROID
CC + A + H(2)O.
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.
CC -!- DOMAIN: THE LEUCINE-RICH HYDROPHOBIC AMINO ACID N-TERMINAL REGION
CC PROBABLY HELPS TO ANCHOR THE PROTEIN TO THE MICROSMAL MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC PIR: A32525; A32525
CC DR PROSITE; PS00086; CYTOCHROME_P450; 1.
CC DR PFAM; PF00067; P450; 1.
CC KW OXIDOREDUCTASE; MONOOXYGENASE; MEMBRANE; HEME; MICROSOME;
CC ENDOPLASMIC RETICULUM; STEROIDOGENESIS; STEROID-BINDING.
CC FT BINDING 427 427
CC FT DOMAIN 341 357
CC FT BINDING 437 437
CC SQ SEQUENCE 492 AA; 55518 MW; 9A721CE0 CRC32;

Query Match 92.7%; Score 38; DB 1; Length 492;
Best Local Similarity 83.3%; Pred. No. 6.87e+00;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 484 GRGESP 489
QY 1 GRGESP 6

RESULT 14
ID CPE1_BOVIN STANDARD: PRT; 495 AA.
AC O18963;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DE CYTOCHROME P450 2EL (EC 1.14.14.1) (CYP11E1).
GN CYP2E1.
OS BUI TARSUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDEA; BOVINA; BOS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HEREFORD; TISSUE-LIVER;
RA VAN RAAK M., NATSUHORI M., LIGTENBERG M., KLEIJ L., TEN BERGHE D.,
RA DE GROENE E.M., VAN MIET A.S., WITKAMP R.F., HORBACH G.J.;
RT "Isolation of a full length cytochrome P450 (CYP2E) cDNA sequence and
RT its functional expression in v79 cells.";
RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
CC ACIDS, AND XENOBIOTICS.
CC -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) - ROH +
CC OXIDIZED FLAVOPROTEIN + H(2)O.
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ001715; E354269; -
CC DR PROSITE; PS00086; CYTOCHROME_P450; 1.
CC DR PFAM; PF00067; P450; 1.
CC KW OXIDOREDUCTASE; MONOOXYGENASE; ELECTRON TRANSPORT; MEMBRANE; HEME;
CC MICROSOME; ENDOPLASMIC RETICULUM.
CC FT BINDING 437 437
CC FT BINDING 437 437
CC SQ SEQUENCE 495 AA; 56827 MW; 3F2B6FCE CRC32;

Query Match 92.7%; Score 38; DB 1; Length 495;
Best Local Similarity 83.3%; Pred. No. 6.87e+00;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 99 GRGESP 104
QY 1 GRGESP 6

RESULT 15
ID YNM7_YEAST STANDARD: PRT; 953 AA.
AC P53917;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 109.8 KD PROTEIN IN CPT1-SFC98 INTERGENIC REGION.
GN YNL127W OR N1221 OR N1875.
OS SACHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACHAROMYCETALES;
OC SACHAROMYCETACEAE; SACHAROMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RX MEDLINE; 96109932.
RA MALLET L., BUSSEREAU F., JACQUET M.;
RT "A 43.5 kb segment of yeast chromosome XIV, which contains MFA2,
RT MEP2, CAP/SRV2, NAM9, FKBI/FPB1/RBP1, MOM22 and CPT1, predicts an
RT adenosine deaminase gene and 14 new open reading frames.";
RL YEAST 11:1195-1209(1995).
RN [2]
RP SEQUENCE OF 26-953 FROM N.A.
RX MEDLINE; 97245296.
RA DE ANTONI A., D'ANGELO M., DAL PERO F., SARTORELLO F., PANDOLFO D.,
RA PALLAVICINI A., LANFRANCHI G., VALLE G.;
RT "The DNA sequence of cosmid 14-13b from chromosome XIV of
RT Saccharomyces cerevisiae reveals an unusually high number of
RT overlapping open reading frames.";
RL YEAST 13:261-266(1997).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z46843; G510324; -
CC DR EMBL; Z71402; E328838; -
CC DR EMBL; Z71403; E239941; -
CC DR EMBL; Z69382; E221820; -
CC DR HSSP; P03036; 3CRO.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 953 AA; 109792 MW; B20E49F6 CRC32;

Query Match 92.7%; Score 38; DB 1; Length 953;
Best Local Similarity 83.3%; Pred. No. 6.87e+00;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 369 GRGESP 374
QY 1 GRGESP 6

Search completed: Thu Dec 23 10:13:31 1999
Job time : 8 secs.

```

**THIS PAGE BLANK (USPTO)**